Supplement

Material and Methods

Patients, healthy controls and cell lines

Probands were enrolled in this study after obtaining written informed consent. The samples were anonymized and all experiments were approved by the Danish Data Protection Agency, the Danish Scientific Ethics Committee, and the Ethical Review Board of the University Hospital, Heinrich-Heine-University, Düsseldorf.

All cell lines were acquired from the German Collection of Microorganisms and Cell Cultures (DSMZ, Brunswick, Germany) and cultured as recommended. The B cell precursor leukemia cell line 697 harbors a t(1;19)(q23;p13.3) translocation coding for *TCF3-PBX1*. Other cell lines used were the *TCF3-PBX1* negative B cell precursor leukemia cell lines REH and HAL-01, the human embryonic kidney cell line HEK-293, the acute myeloid leukemia cell line HL-60, and the acute monocytic leukemia cell lines THP-1 and MV4-11.

Processing of umbilical cord blood samples

Umbilical cord blood samples from 340 healthy neonates were collected in blood collection tubes with EDTA as anticoagulation additive. All samples were processed within 24 h (median 12 h) from birth. Mononuclear cells (MNCs) were separated by Ficoll density centrifugation, washed with RPMI 1640 medium and resuspended in 2 ml RPMI 1640 medium containing 10% BSA. MNCs were transferred to cryovials and a cryopreservative solution was added 1+1. The vials were frozen at -80°C in a 5100 Cryo 1C Freezing Container and subsequently transferred to liquid nitrogen.

Enrichment of CD19 positive cells

CD19 positive B cells were enriched from at least 3×10^6 cells using magnetic Dynabeads CD19 pan B (Invitrogen, Carlsbad, CA, USA), as recommended by the manufacturer. Briefly, MNCs derived from cord blood were thawed, diluted 1:1 with MACS buffer (Miltenyi, Bergisch

Gladbach, Germany) and centrifuged at 600 x g and 4°C for 10 min. Cell pellets were resuspended in 1 ml MACS buffer. 50 µl magnetic beads coated with CD19 antibodies sufficient for binding of 2.5 x 10^7 $CD19^+$ cells were washed once with MACS buffer and added to each sample. Cells and beads were incubated at 4°C with gentle tilting for 20 min. Subsequently, cells were washed with MACS buffer and separated using a magnet twice. Cells were resuspended in 200 µl 1x PBS. Purity of the $CD19^+$ cell fraction was tested by flow cytometry to be above 95%.

DNA isolation

Genomic DNA was isolated form $CD19^+$ cells and cell lines employing the QIAamp Blood Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. DNA derived from $CD19^+$ cells was finally eluted in 80 µl AE buffer.

GIPFEL screening

The GIPFEL procedure was carried out basically as described.[1] Briefly, genomic DNA was digested enzymatically with *MfeI* (NEB, Ipswich, MA, USA) to achieve defined fragmentation. For the umbilical cord blood screening process, the entire vial of cells, a maximum of 1.8 ml, was subjected to the GIPFEL process. The number of B cells was estimated to range between 2.5 x 10^5 and 3 x 10^6.

After the restriction enzyme digest, the DNA was purified using the QIAquick Gel Extraction Kit (QIAGEN) with a final DNA incubation and elution in 50 µl preheated water at 60°C for 5 min. The purified DNA was ligated with T4 DNA ligase (NEB) at 24°C for 2 h to achieve circularization. To remove residual linear DNA 100 U of Exonuclease III (NEB) were added and incubated at 37°C for 30 min. Then exonuclease and ligase were heat inactivated at 95°C for 5 min. To purify and concentrate the DNA, an ethanol precipitation was carried out and the dried DNA pellet was resuspended in 25 µl nuclease free water. PCR was used to detect the ligation joints produced by circularization. To this end, the DNA was pre-amplified by 25 cycles of PCR with multiplexed primers using the GoTaq qPCR Master Mix (Promega, Mannheim,

Germany) and a T-Gradient Thermoblock PCR cycler (Biometra, Göttingen, Germany). One reverse primer covered the breakpoint cluster region (BCR) of *TCF3*, while 36 forward primers covered the BCR of *PBX1*. Forward primers were pooled in four groups of seven primers each and one group of eight. Two additional primers amplified a product outside of the *PBX1* BCR, which served as a positive control.

The PCR products were used as templates for a Real-Time PCR performed on a CFX Real-Time PCR Detection System (BioRad, Hercules, CA, USA) using the same master mix with the same forward primer bundles, but the reverse primer was substituted by a nested reverse primer (Supplemental Figure S3A).

When a sample yielded a positive result, a new Real-Time PCR was done, this time with demultiplexed forward primers. If one of the forward primers still produced a positive result, the PCR was repeated with this forward primer to analyze the product on an agarose gel and Sanger sequenced if a specific DNA band was still present (Supplemental Figure S3B).

Validation of GIPFEL results by identification of patient-specific genomic TCF3-PBX1 breakpoints

MfeI-digested and circularized DNA was used as a template for PCR amplification. Briefly, a forward primer located at the 3' end of *TCF3* exon 16 was combined with a *PBX1* reverse primer close to the identified ligation joint but facing towards the breakpoint. The PCR reaction containing 2 µl template DNA, primers (5 µl each of 10 µM forward and reverse primers, Supplemental Table S2), 25 µl 2x Phusion HF Master Mix (Thermo Fisher, Waltham, MA, USA), and 13 µl water was mixed and the PCR was carried out on a T-Gradient Thermoblock PCR cycler (Biometra, Göttingen, Germany). Cycling conditions were as follows: 98°C for 5 min, 35 cycles of 98°C for 30 s, 65°C for 30 s, and 72°C for 5 min, followed by a final elongation at 72°C for 10 min. The PCR products were loaded and separated on an agarose gel, cut from the gel, and purified with the QIAquick Gel Extraction Kit (QIAGEN, Hilden, Germany), according to manufacturer's recommendations. Subsequently, the samples were Sanger sequenced (Supplemental Figure S3C). In brief, DNA was mixed with 4 µl Big Dye Terminator (Applied Biosystems, Waltham, MA, USA), 0.5 µl 10 µM sequencing primer

(Supplemental Table S2), and water (20 µl final volume). The samples were then PCR amplified with 26 cycles of 96°C for 30 s, 55°C for 30 s, and 60°C for 4 min. Subsequently, the sequencing was carried out on an ABI 3130 sequencer according to the recommendations of the manufacturer (Applied Biosystems). For N15, the sequencing had to be repeated with new primers that lay closer to the breakpoint (Supplemental Table S2).

Supplemental Tables

Supplemental Table S1: Validation of GIPFEL screening of 7 human cell lines by RT-qPCR.

Cell line	Present Translocation	Origin	Translocation t(1;19)	GIPFEL
697	t(1;19)	B-cell precursor leukemia	Positive	positive
REH	t(12;21)	B-cell precursor leukemia	negative	negative
HAL-01	t(17;19)	B-cell precursor leukemia	negative	negative
HL60	t(5;17), t(9;14), t(16;17)	Acute myeloid leukemia	negative	negative
MV4-11	t(4;11)	Acute monocytic leukemia	negative	negative
THP-1	t(9;11)	Acute monocytic leukemia	negative	negative
HEK293	t(5;?)	Embryonal kidney cells	negative	negative

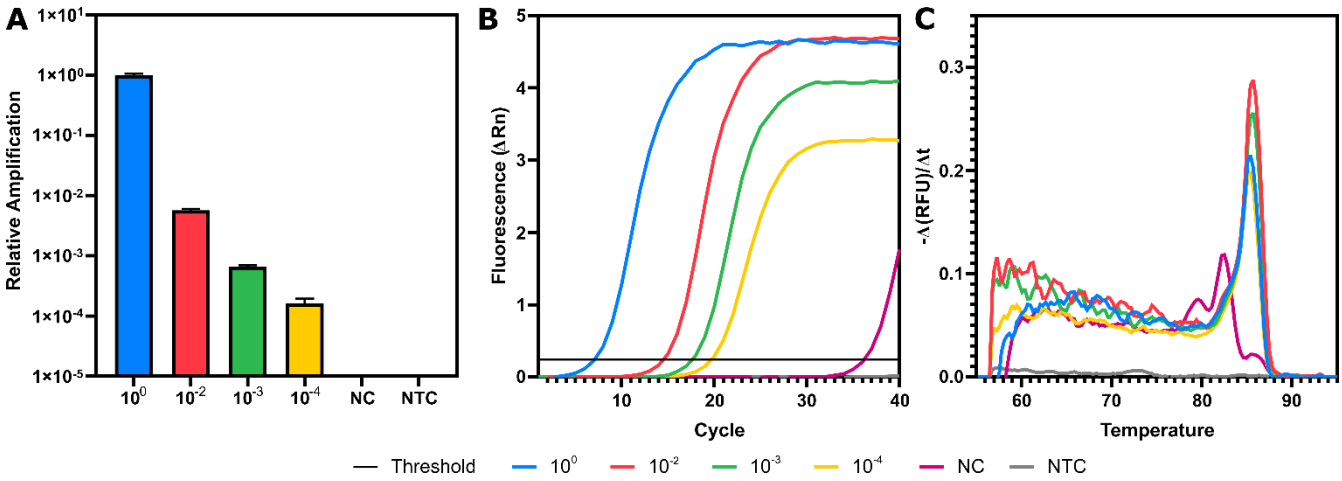
Supplemental Table S2: Primers for PCR amplification of DNA circles and for Sanger sequencing of chromosomal breakpoints.

Name	Sequence [5'-3']
TCF3-BP-f	GTCTCGGCCTCCCGACTCCTACAGTG
TCF3-BP-f2	CACCAGCCCAGGAATCCTGCCTG
PBX1-BP8-r	TCAGCGTGGGATTTCAGAGCTTTACCTTGAGGATG
PBX1-BP20-r	AAGGCGCCAACACACTTCAACATACTTTGCTAATATAATTATTTGC
PBX1-BP20-r2	GTAGGCACGCTTCCATTTGCTTAAGTC

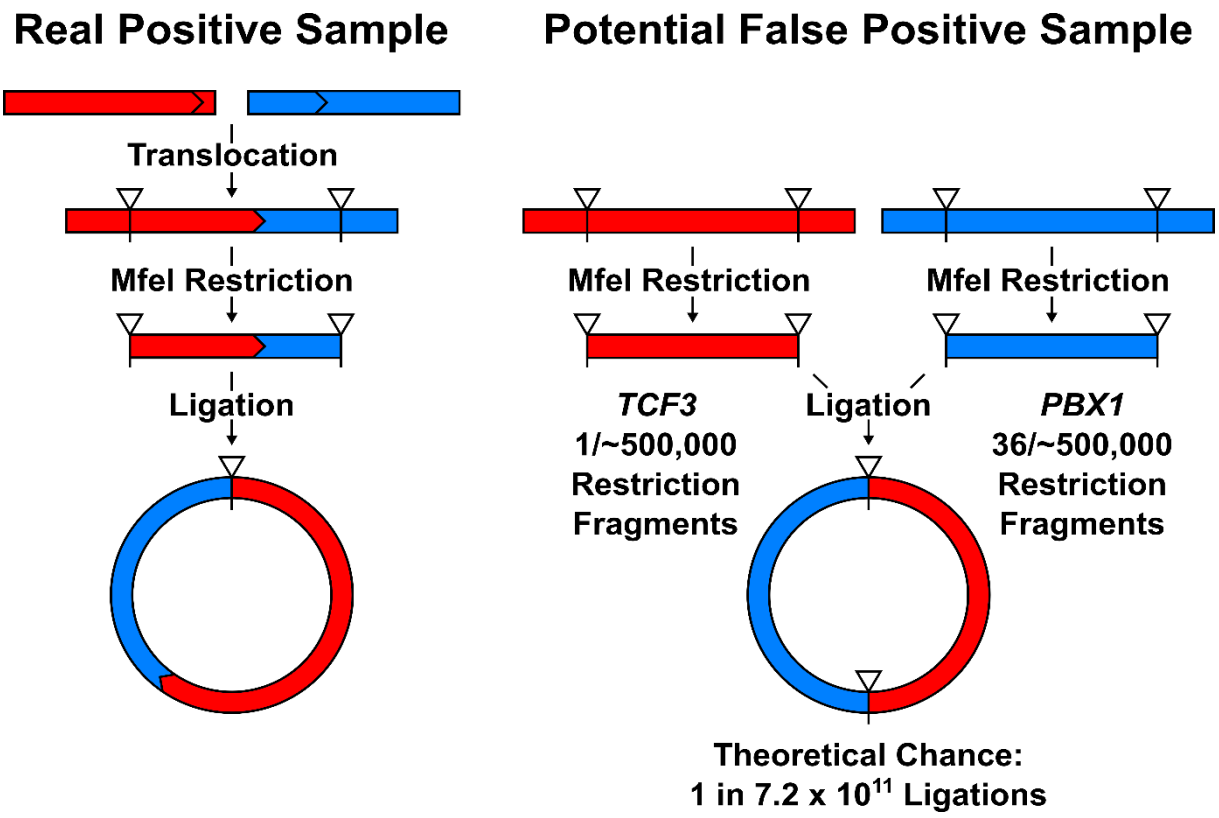
Supplemental Table S3: Reported Breakpoints of *TCF3-PBX1* fusions. The exact position of the break is given (GRCh38) as well as the 20 bp 5' and 3' of the breakpoint. Inserted nucleotides are also shown. Colors indicate identical breakpoint positions. Samples in italics indicate cell lines.

See file *Supplemental Table S3.xlsx* for table.

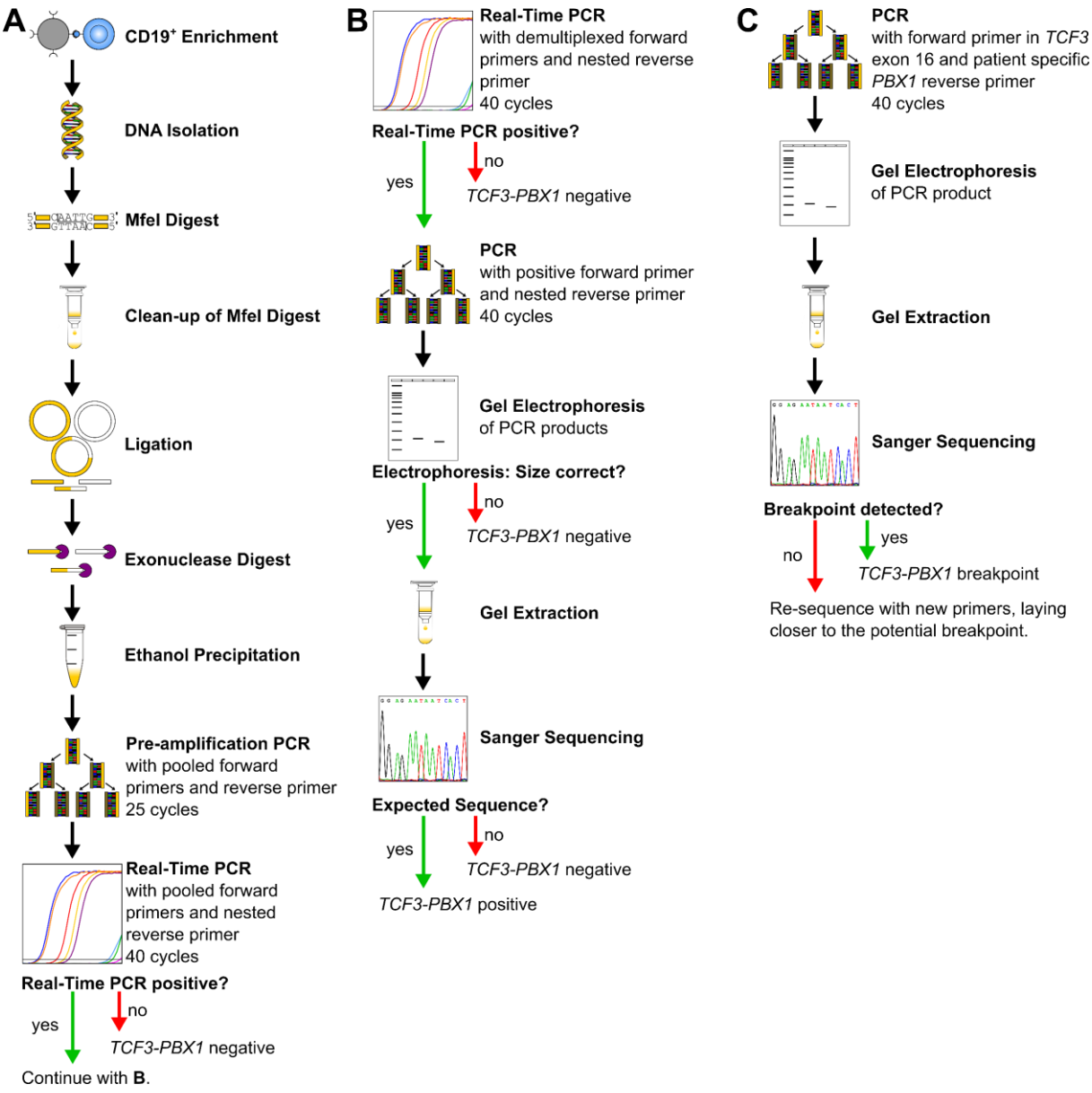
Supplemental Figures



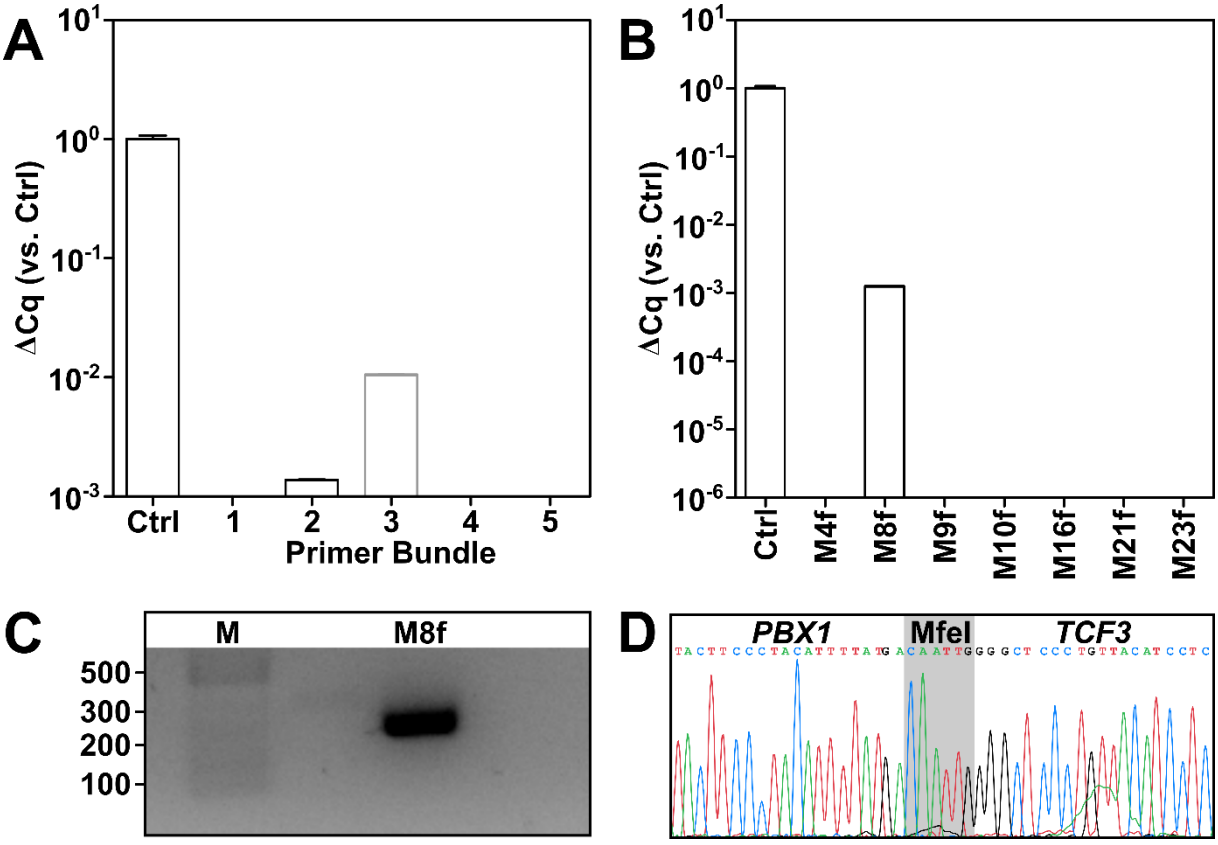
Supplemental Figure S1: Sensitivity of the GIPFEL method for *TCF3-PBX1*. Results of the Real-Time PCR of a dilution series of *TCF3-PBX1*$^{+}$ cell line 697. **(A)** Amplification of the dilutions relative to undiluted 697 (10^{0}). **(B)** Amplification plot of the dilution series. **(C)** The melt curve of the dilution series shows the same product for all tested dilutions. NC = negative control, NTC = non-template control. Data taken from[1].



Supplemental Figure S2: Model explaining the high specificity of GIPFEL. *Left panel:* A true positive sample. The translocation generates a *Mfe*I restriction fragment that contains material from chromosome 1 and 19. After restriction and ligation, circular DNA allows the detection of the ligation joint through the GIPFEL procedure. *Right panel:* An assumed false positive signal. DNA without the investigated translocation is *Mfe*I restricted and the restriction fragments are ligated to form circular DNA without a true breakpoint. When the human genome is digested with *Mfe*I, ≈500,000 fragments are generated. Only one of those covers the *TCF3* breakpoint cluster region and can be detected by GIPFEL. For *PBX1*, 36 fragments can be detected. The probability that these fragments are ligated together by chance is therefore very low. As these fragments have to be ligated in the correct orientation for the primers to generate a PCR product, the probability has to be divided by two. Thus, the theoretical probability that a false positive signal occurs is $\approx 7.2 \times 10^{-11}$. A false positive result in addition requires two independent ligation events.



Supplemental Figure S3: The workflow of cord blood screening employing a modified GIPFEL technique is presented. (A) CD19⁺ B cells were enriched from newborn cord blood. DNA was isolated, fragmented using *Mfe*I and purified. The DNA was ligated to achieve circularization and residual linear DNA was digested. DNA circles were purified and PCR (preamplification) and Real-Time PCR were used to detect the ligation joints produced by circularization. The protocol was carried out with each sample**. (B)** Optional continuation protocol that was carried out if the Real-Time PCR in (A) generated a putative positive result. Then a new Real-Time PCR was carried out with demultiplexed forward primers. If one of the forward primers produced a positive result, the PCR was repeated with this forward primer and the reverse primer. The products were then analyzed on an agarose gel and Sanger sequenced. In case of negative results, no further validation steps were done. **(C)** Further development of the original GIPFEL technique allowed the identification of patient-specific breakpoints with base pair resolution. To this end, circularized DNA was PCR amplified using specific primers that hybridized to *TCF3* exon 16 and the known *PBX1* ligation joint region and faced towards the unknown breakpoint. The PCR products were separated on an agarose gel, cut from it, and purified. Eventually, the products were Sanger sequences. When the breakpoint was not yet reached, the sequencing was repeated with new primers that lay closer to the breakpoint.



Supplemental Figure S4: Cord blood sample N141 harbors a *TCF3-PBX1* fusion as revealed by DNA-based GIPFEL screening. (A) Result of the first Real-Time PCR. The amplification of the five primer bundles (1 through 5) was compared to the amplification of the internal *PBX1* wild-type control (Ctrl). Primer bundle 2 exceeds the detection limit of 10^{-4}; primer bundle 3 (in grey) produces primer dimers, as identified by the melt curves (not shown). **(B)** Amplification plot of the second Real-Time PCR. The forward primers of bundle 2 were demultiplexed and primer PBX1-M8f was identified as the one responsible for the amplification. **(C)** Agarose gel after a PCR with PBX1-M8f and the reverse primer TCF3-M1r-n. The PCR shows the expected product of 273 bp. **(D)** Sanger sequencing result of cord blood N141. The expected sequences flanking the *MfeI* ligation joint were identified, indicating the fusion of *PBX1* segment M8 to *TCF3* segment M1.

References

1. Füller E, Schäfer D, Fischer U, et al. Genomic Inverse PCR for Exploration of Ligated Breakpoints (GIPFEL), a New Method to Detect Translocations in Leukemia. *PLoS One*. 2014;9(8):e104419.